

TECHNICAL SUPPLEMENT 3: MAPPING SOURCE-SPECIFIC HOSPITAL IDENTIFIERS TO AHA HOSPITAL IDENTIFIERS

TABLE OF CONTENTS

INTRODUCTION	2
RECONCILING HOSPITAL IDENTIFIERS	2
Electronic Linkage of Source and AHA Hospital Identifiers	2
Resolution of Unmatched Hospitals	4
RULES FOR RESOLVING PROBLEM HOSPITALS	5
Openings	5
Closures	6
Mergers	6
Demergers	7
Changes in Hospital Characteristics	7

INTRODUCTION

The American Hospital Association (AHA) definition of "community hospital" was used to select hospitals for the HCUP databases. Therefore, for each participating data source and for each year, it was necessary to reconcile the data source's identification of the hospital with the identification of the hospital in the associated AHA Annual Survey. The list of all such linkages is called a *crosswalk*.

Once these linkages were established, data from the AHA Annual Survey were used to:

- identify facilities that are defined by the AHA as "community hospitals;"
- classify each community hospital into a stratum for sampling for the Nationwide Inpatient Sample (NIS);
- add various types of information about the hospital (such as its county FIPS code) to the inpatient records; and
- identify community hospitals listed in the AHA Annual Survey for which no inpatient data were supplied by the data source.

This Technical Supplement addresses the procedures used to identify the appropriate linkages between AHA hospital identifiers and hospitals represented in the inpatient data supplied by each data source.

Note: In this document, *data source* always refers to the source of the inpatient data.

RECONCILING HOSPITAL IDENTIFIERS

The goal is to identify an appropriate AHA hospital identifier for each source hospital.

To begin, relevant data elements are extracted from inpatient data and from the AHA data for each year. The two elements extracted from the inpatient data are:

- the source-specific hospital identifier, and
- a count of the hospital's inpatient records for each quarter and for the year.

Electronic Linkage of Source and AHA Hospital Identifiers

First, the hospital identifiers used by the data source (hereafter referred to as DSHOSPID) are linked electronically to the relevant data elements extracted from the AHA Annual Survey data. AHA identifiers include all hospitals in the state, not just the community hospitals that are eligible for inclusion in HCUP.

A SAS merge step is used to link the DSHOSPIDs to AHA identifiers. The specific variables used in the merge depend on the information provided by the data source. In order of preference, these variables are:

- hospital name, city, and zip code;
- hospital name; or
- any other unique variable that is available – e.g., Medicare provider number.

The AHA and the data source often use different methods to represent components of a hospital's name (e.g., "Community General Hospital" may be represented as "Community General Hosp" by the AHA and as "Community Gen Hosp" by the data source). Hence, before the SAS merge, the AHA and the source's hospital names are transformed into a uniform link variable, which imposes similar methods of abbreviations, lowercase and uppercase letters, and different characters. This effectively reduces the number of non-merges that occur simply because of different methods of representing the hospital name's components.

Three types of linkages result from this step, as shown in Table 1.

Table 1. Linkages Between AHA and Data Source Identifiers

Row	AHA Identifier	DSHOSPID	Link?
1	present	present	yes
2	present	absent	no
3	absent	present	no

Approximately 80 percent of DSHOSPIDs link to AHA identifiers in this step. (These successful links are represented by Row 1 in Table 1). The other 20 percent of DSHOSPIDs (Rows 2 and 3) must be linked manually, using the process described below.

Prior experience has shown that a large majority of the hospitals failing to link in this step will usually be of the following types:

- closures,
- openings (new hospitals),
- mergers,
- demergers,
- dates of changes in the hospital structure that differ between the data source and the AHA Annual Survey, and
- levels of aggregation that differ between the data source and the AHA Annual Survey (e.g., the data source treats two separate facilities as two hospitals, while the AHA Annual Survey treats the two facilities as a single hospital, or vice versa).

Resolution of Unmatched Hospitals

The goal in this step is to identify an appropriate AHA hospital identifier for each source hospital that was not matched electronically above.

Several external sources of information are used to reconcile the unmatched source hospitals (Row 3) and the unmatched AHA hospitals (Row 2). These are:

- *Source Documentation:* This information, received from the data source, usually contains a list of hospitals, their cities (specific addresses are not always included), and the source's hospital identifier. This documentation is the primary source for finding missing information (e.g., specific names and addresses) for an unmatched hospital.
- *AHA Summaries:* The AHA Summary of Registered Hospitals and the AHA Summary of Nonregistered Hospitals, which are usually delivered with the annual *AHA Guide*, document additions and deletions reflected in the hospitals' responses to the AHA Annual Survey.
- *AHA Guide:* The *AHA Guide*, an annual hard-copy volume published by the AHA, provides a wealth of information about registered hospitals. The *AHA Guide* includes an entire section on multihospital health-care systems that identifies the hospitals included in specific multihospital systems. The *AHA Guide* also provides information about individual hospitals (organized by state, and within each state, by city), which includes:
 - Information also available from the AHA Annual Survey data; for example:
 - type of service (general medical/surgical, rehabilitation);
 - average lengths of stay (long- or short-term);
 - type of ownership; and
 - numbers of beds, admissions, births, etc.
 - Information about hospitals embedded within the organizational structure of another hospital; for example:
 - Binghamton – Broome County, NY*
United Health Services (includes Binghamton General Hospital, Mitchell Ave. ...; Wilson Memorial Hospital, ... Harrison St. ...)
 - Information about changes in a hospital's name; for example:
 - Dobbs Ferry – Westchester County, NY*
Community Hospital at Dobbs Ferry (formerly Dobbs Ferry Hospital)
 - References to a new hospital name or new location; for example:
 - Delhi – Delaware County, NY*
A. Lindsay and Olive B. O'Connor Hospital. See Mary Imogene Bassett Hospital, Cooperstown.

- *Record counts generated from the supplied inpatient data:*
The number of discharges reported in the inpatient data is compared to the number of discharges reported to the AHA. While this information is rarely definitive in linking source identifiers to AHA identifiers, it is sometimes useful in identifying a link to an AHA hospital, and provides a means of validating linkages obtained by other means. This information is especially useful in distinguishing a single hospital from two combined hospitals.

When it is not clear what needs to be done to a hospital or group of hospitals, an AHRQ analyst is consulted.

The reconciliation process is complete when the following facts are confirmed:

- all source hospital identifiers have been assigned an HCUP hospital identifier (HOSPID), and
- all hospitals (HOSPID) have only one FIPS county code assigned.

Hospitals composed of multiple facilities in different locations are assigned the FIPS county code of the major facility, as defined by the AHA.

RULES FOR RESOLVING PROBLEM HOSPITALS

Following are the rules used for resolving problem hospitals. In these examples, the HCUP hospital identifier (HOSPID) starts with "SS" to indicate the state FIPS code:

The HCUP hospital identifier (HOSPID) reflects the AHA view of a hospital and is a randomly assigned number based on the AHA hospital identifier (IDNUMBER). If the data source reports the data from facilities that are combined in the AHA hospital definition, the IDNUMBER and the HOSPID will be the same for all the facilities. In the following example, three different source identifiers are considered to be part of one facility as defined by the AHA:

<i>Year</i>	<i>Data Source</i>	<i>AHA IDNUMBER</i>	<i>HOSPID</i>
1990	165 (Acute Care Unit)	910140	SS089
1990	165P (Psychiatric Unit)	910140	SS089
1990	166S (Swing Bed Unit)	910140	SS089

Openings

The AHA IDNUMBER and HCUP HOSPID are assigned to a newly opened hospital only when the hospital has first been recognized by the AHA for a particular survey year, even if the data source supplies data for an earlier time period. For example, the data source supplied data for a hospital starting in 1989, but the AHA first recognized the hospital in 1991:

<i>Year</i>	<i>Data Source</i>	<i>AHA IDNUMBER</i>	<i>HOSPID</i>
1989	86-0601625		
1990	86-0601625		
1991	86-0601625	860001	SS014
1992	86-0601625	860001	SS014

Closures

When a hospital closes (in the AHA's view), the AHA IDNUMBER and HCUP HOSPID are carried forward if there are inpatient data available from the data source. In this example, the AHA considered the hospital closed in 1990, but the data source still supplied data:

<i>Year</i>	<i>Data Source</i>	<i>AHA IDNUMBER</i>	<i>HOSPID</i>
1988	047	450520	SS171
1989	047	450520	SS171
1990	047	450520 (closed)	SS171

Mergers

When two or more hospitals merge (in the AHA's view), the IDNUMBER (along with the HOSPID) of the merged entity is assigned to all its component hospitals even if they continue reporting separately to the state. In this example, two hospitals have different source identifiers, but starting in 1990 are considered one facility by the AHA because of a merger:

<i>Year</i>	<i>Data Source</i>	<i>AHA IDNUMBER</i>	<i>HOSPID</i>
1989	036	450400	SS091
1990	036	450002 (merger)	SS013
1991	036	450002 (merger)	SS013
1992	036	450002 (merger)	SS013
1989	126	451750	SS169
1990	126	450002 (merger)	SS013
1991	126	450002 (merger)	SS013
1992	126	450002 (merger)	SS013

Demergers

When hospitals demerge (in the AHA's view), the component hospitals are assigned a new AHA IDNUMBER or the one they previously had. The HCUP HOSPID follows the AHA IDNUMBER, so that the HOSPID changes if the IDNUMBER changes and the HOSPID is reused if the IDNUMBER is reused. In this example, a hospital demerges in 1989 into two facilities:

<i>Year</i>	<i>Data Source</i>	<i>AHA IDNUMBER</i>	<i>HOSPID</i>
1988	562	220515 (merger)	SS051
1989	562	220547 (demerger)	SS026
1990	562	220547 (demerger)	SS026
1988	561	220515 (merger)	SS051
1989	561	221240 (demerger)	SS037
1990	561	221240 (demerger)	SS037

Changes in Hospital Characteristics

(Note: The following decision is made only after AHRQ is consulted.) If during HCUP processing of the inpatient data, summary statistics on the distribution of length of stay look questionable for a community hospital (e.g., the mean length of stay is considerably greater than 30 days), the AHA community flag is investigated. If the AHA community flag was imputed from previous years because a hospital did not report to the AHA – and the data source can confirm that the facility is no longer a community hospital – the AHA identifier is still assigned to the facility, the community flag is imputed, and the hospital is excluded from the HCUP inpatient databases. In this example, the facility was considered a non-community hospital starting in 1990:

<i>Year</i>	<i>Data Source</i>	<i>AHA IDNUMBER</i>	<i>HOSPID</i>	<i>Community Flag</i>
1988	86-0201864	860575	SS090	1
1989	86-0201864	860575	SS090	1
1990	86-0201864	860575	SS101	0 (changed)*
1991	86-0201864	860575	SS101	0 (changed)*
*Hospitals with the community flag indicator of "0" are not processed.				